Statistics

Lecture 22

Feb 19-8:47 AM

Find minimum Sample Size needed if we wish to construct conf. interval for Population proportion with error not to exceed 8%..

$n = \hat{p}\hat{q}\left(\frac{Z_{\alpha/2}}{E}\right)^2$

Always Round-up

n NO C-level → use 95%

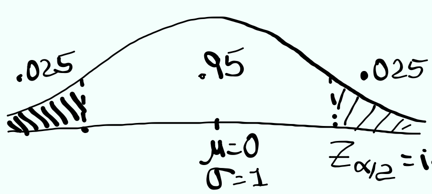
$E = .08$

NO $\hat{p}$ & $\hat{q}$

$n = (.5)(.5)\left(\frac{1.960}{.08}\right)^2$

$n = 150.0625$

$\boxed{n = 151}$



$Z_{\alpha/2} = \text{invNorm}(.975, 0, 1) \approx 1.960$

May 7-1:50 PM

In a Sample of 20 apartments, the mean rent was \$2500/mo. with standard deviation of \$375/mo.

$n = 20$
$\bar{x} = 2500$
$S = 375$

C-level: .99

Find 99% Conf. interval for the mean rent of all apartments. $< \mu <$

σ Known → Z Interval

σ unKnown → T Interval

STAT TESTS TInterval

inpt: Stats

$df = n-1$
$= 19$

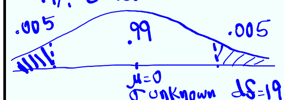
$\bar{x} = 2500$
$S = 375$
$n = 20$
C-level: .99

$\boxed{2260 < \mu < 2740}$

$E = \frac{2740 - 2260}{2} = \boxed{240}$

$\bar{x} = \frac{2740 + 2260}{2} = \boxed{2500}$

Find $t_{\alpha/2}$ for 99% C-level



$t_{\alpha/2} = \text{invT}(.995, 19) = \boxed{2.861}$

$E = t_{\alpha/2} \cdot \frac{S}{\sqrt{n}} = 2.861 \cdot \frac{375}{\sqrt{20}} = 239.902 \approx \boxed{240}$

May 7-1:56 PM

I randomly selected 15 days in LA area. Here are the temp. :

65	72	80	92
58	75	88	95
100	68	78	94
90	85	79	

Store in L1

Find

1) $\bar{x} \approx 81$
2) $S \approx 12$

} Round to whole #

3) $S^2 = \frac{15577}{105}$

} Reduced fraction

Find 90% Conf. interval for mean temp. of all days in LA.

$\boxed{76 < \mu < 86}$

σ Known → Z Interval

σ unKnown → T Interval

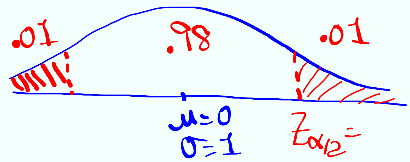
$\bar{x} = \frac{86 + 76}{2} = 81$

$E = \frac{86 - 76}{2} = 5$

May 7-2:09 PM

Find minimum Sample Size needed if we wish to construct $\boxed{98\%\ \text{Conf. interval}}$ for the mean temp. of all days and error not to exceed 4 degrees.

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2$$



Always round-up

invNorm(.99,0,1) = 2.326

when σ unknown → use S

$$n = \left(\frac{z_{\alpha/2} \cdot S}{E}\right)^2 = \left(\frac{2.326 \cdot 12}{4}\right)^2 = 48.692\ldots = \boxed{49}$$

May 7-2:17 PM

SG 23

Testing Claims:

A claim is made and we want to know if it is a valid claim or invalid claim.

we wish to determine its validity.

If claim is valid ⇒ we Support it.
Fail-to-Reject

If claim is invalid ⇒ we reject it.

Claim could be about

1) Population Proportion p
2) Population Mean μ
3) Population Standard deviation σ

May 7-2:24 PM

College claims $\underbrace{\text{10\% of all students}}_{}$ smoke

$p = .1$ claim

I claim the mean age of all students is below 35.

$\mu < 35$ claim

Math department claims that standard deviation of all exam scores is at least 12.

$\sigma \geq 12$ claim

May 7-2:30 PM

Common Sense

Valid claim $\Longleftrightarrow$ Fail-to-Reject

Invalid claim $\Longleftrightarrow$ Reject

Possible Errors:

Valid claim but we reject it.

Invalid claim but we fail-to reject.

May 7-2:34 PM

why are we doing this?

to determine claim's validity

Testing Methods:

1) Traditional Method
2) P-Value Method
3) Confidence Interval Method

Regardless of the method, the final Conclusion must be the Same.

Conclusion \ claim	VALID	Invalid
Support	Good Decision	Error
Reject	Error	Good Decision

May 7-2:38 PM

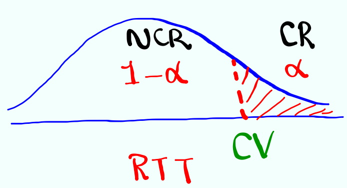
Testing Types:

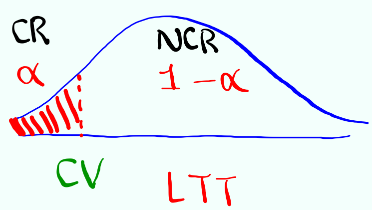
1) Right-Tail Test **RTT**
2) Left-Tail Test **LTT**
3) Two-Tail Test **TTT**

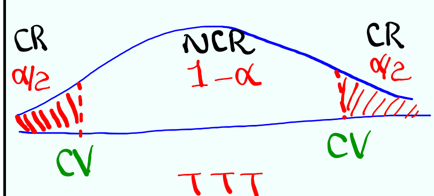
with every testing, there is a Significance level α.

$0 < \alpha < 1$

If α not given use .05







$\alpha \rightarrow$ Critical Region

$1-\alpha \rightarrow$ Non-Critical Regio

May 7-2:47 PM

Testing Process:

1) Set-up H_0 & H_1

H_0: Null Hypothesis

H_1: Alternative Hypothesis, H_a

2) Find all Critical Values

Drawing, labeling, shading, and full TI Command required.

3) Find Computed Test Statistic CTS and P-Value P.

Formula or full TI Command required.

4) Use $\boxed{\text{testing chart}}$ to determine the validity of H_0 & H_1.

5) Draw final Conclusion about the claim.

claim could be H_0 or H_1.

Final Conclusion

Reject the claim OR $\underbrace{\text{FTR}}_{\text{Fail-to-Reject}}$ the claim

May 7-2:56 PM

More on H_0 & H_1:

H_0 must contain equal sign $=, \geq, \leq$

H_1 cannot contain equal sign $\neq, <, >$

Keywords for H_0:

is, equal, same, at least, at most, not different

Keywords for H_1:

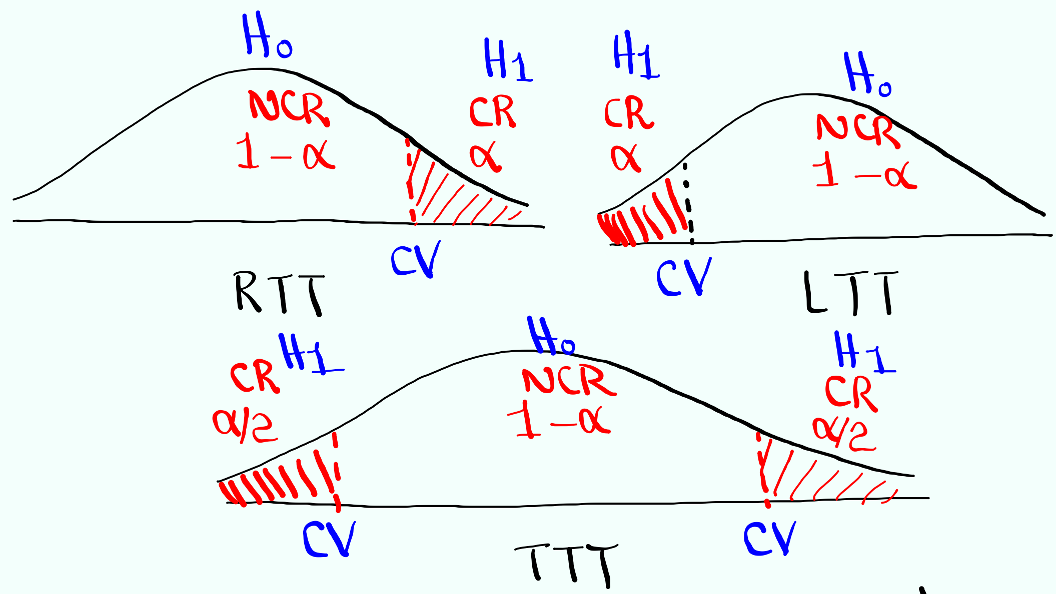
is not, not equal, not same, more than, less than, different, below, above, grear than, exceed, ...

H_0: $=$	H_0: $\geq$	H_0: $\leq$
H_1: $\neq$	H_1: $<$	H_1: $>$
TTT	LTT	RTT

H_1 helps us to determine testing types.

Always identify claim and Testing Type.

May 7-3:04 PM



$P(H_0 \text{ valid}) = 1 - \alpha = P(H_1 \text{ invalid})$

$P(H_0 \text{ invalid}) = \alpha = P(H_1 \text{ valid})$

May 7-3:13 PM

College claims $\underbrace{10\% \text{ of all}}_{}$ students Smoke

$P = .1$
↑
H_0

$H_0: P = .1$ claim

$H_1: P \neq .1$ TTT

I claim the mean age of all students is below 32 Yrs.

$\mu < 32$
↑
H_1

$H_0: \mu \geq 32$

$H_1: \mu < 32$ claim, LTT

Department claims standard deviation of all exam Scores is at most 12.

$\sigma \leq 12$
↑
H_0

$H_0: \sigma \leq 12$ claim

$H_1: \sigma > 12$ RTT

May 7-3:19 PM

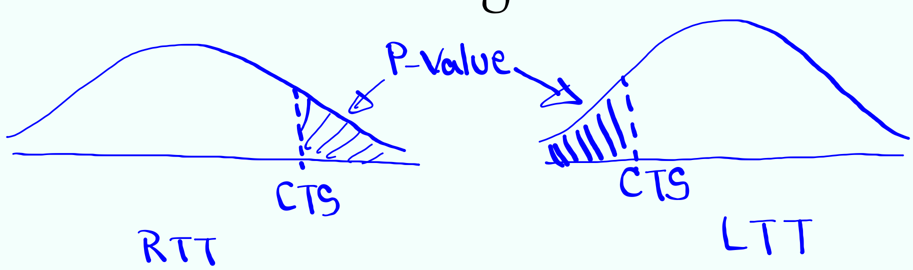
Type I & II errors for H_0:

Action \ Reality	H_0 Valid	H_0 invalid
Support H_0	Good Decision	Type II Error
Reject H_0	Type I error	Good Decision

May 7-3:27 PM

CTS Computed Test Statistic

P-Value is the area of the tail marked by CTS.

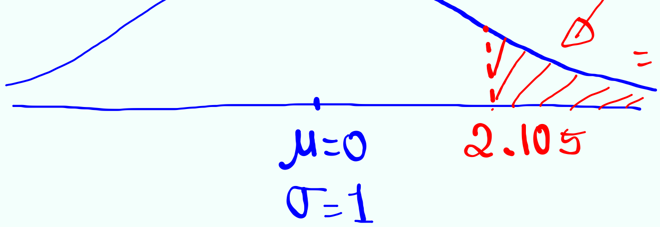


If the Process is TTT,

multiply the tail area by 2

May 7-3:32 PM

find P_Value if CTS $Z=2.105$ and RTT.



Area = P_Value

$= \text{normalcdf}(2.105, E99, 0, 1)$

$= \boxed{.018}$

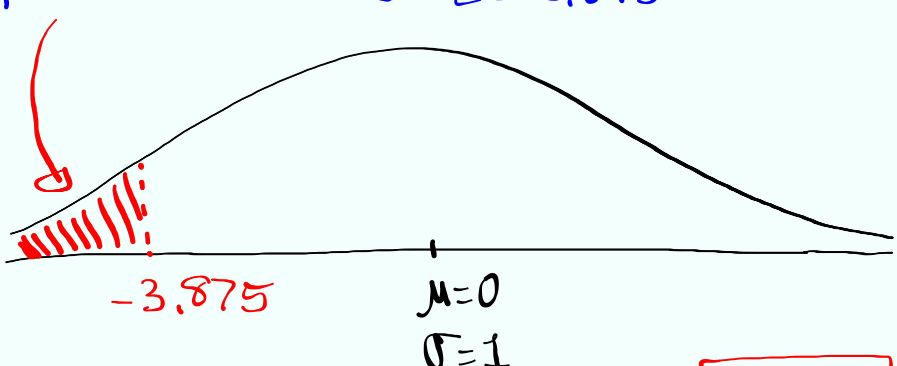
If it was TTT

$\text{P-value} = 2(.018)$

$= \boxed{.036}$

May 7-3:35 PM

find P_Value if CTS $Z=-3.875$ and LTT.



$\text{normalcdf}(-E99, -3.875, 0, 1) = \boxed{5.3 \times 10^{-5}}$

if it was TTT

$\text{P-value} = 2(5.3 \times 10^{-5}) = \boxed{1.1 \times 10^{-4}}$

May 7-3:38 PM